Supplementary Tables

Supplementary Table 1 | Summary of statistical results from ANOVA analyses.

Mean quantity ± s.e.m.	Test	Days	*P*-value	Deg. freedom	Error deg. freedom	Total deg. freedom	χ^2
152 ± 14 active cells per day per mouse (**Extended Data Fig. 5g**)	One-way Friedman ANOVA	6	0.31	$df_{days} = 5$	$df_{err} = 55$	$df_{total} = 71$	5.9
49 ± 2% of cells active per day, in the sets of all cells, CS^+-responsive cells, and CS^--responsive cells (**Extended Data Fig. 5f**)	Two-way Friedman ANOVA	6	>0.05 for all 3 *P*-values	$df_{days} = 5$ $df_{group} = 2$ $df_{interaction} = 10$	$df_{err} = 198$	df_{total} = 215	1.6–7.5
$\Delta_4 = -2 \pm 2\%$; PVD between CS^+ during extinction sessions and the mean CS^+ before conditioning (**Fig. 4e**)	One-way Friedman ANOVA	3	0.37	$df_{CS+} = 11$	$df_{err} = 121$	df_{total} = 143	12
Measured quantity	**Test**	**Groups**	***P*-value**	**Deg. freedom**	**Error deg. freedom**	**Total deg. freedom**	χ^2
Locomotor parameters: Distance traveled, speed, acceleration of mice with zero, one, or bilateral implants in BLA (**Extended Data Fig. 3b**)	One-way Kruskal-Wallis ANOVA	3	≥ 0.05 for all 3 locomotor parameters	$df_{group} = 2$	$df_{err} = 31$	$df_{total} = 33$	10–12
% of time spent freezing by mice with zero, one, or bilateral implants in BLA, or with bilateral implants and muscimol injection (**Extended Data Fig. 3c,d**)	One-way Kruskal-Wallis ANOVA	4	≥ 0.05 for CS^+ and CS^- tones	$df_{group} = 3$	$df_{err} = 42$	$df_{total} = 45$	10–12

Supplementary Table 2 | An example tri-conditional rule learning that is consistent with the BLA ensemble neural activity data and an ensemble level model of supervised learning.

CS input	US input	Global signal accompanying US	Plasticity
Yes	No	No	CS-evoked responses are stable
Yes	No	Yes	CS-evoked responses depress
Yes	Yes	Yes	CS-evoked responses potentiate

Supplemental Note

Can a bi-conditional cellular learning rule account for the plasticity of BLA ensemble neural coding during associative fear conditioning?

The traditional cellular, Hebbian hypothesis of fear learning invokes a bi-conditional learning rule and posits that among the cells receiving CS^+-related inputs, those activated by the US will potentiate their responses to the CS^+ (Ref. 1). This longstanding hypothesis gives an account of the behavioral conditioning but does not fit well with the data (**Figs. 2**, **3e**,**f** and **Extended Data Fig. 6**) concerning the diverse plasticity of CS^+-evoked responses of BLA neural ensembles during and after associative fear conditioning. The following set of observations in BLA is challenging to explain with Hebbian potentiation, or with any hypothesis based on a bi-conditional learning rule:

- During learning, there was up- and down-regulation of neurons' CS^+- and US-evoked responses, demonstrating that single neuron plasticity is diverse and bi-directional during associative fear learning (**Figs. 2d, 3e**).
- Most of the neurons that potentiated their CS^+-evoked responses during learning did not respond to the US (**Fig. 3f** and **Extended Data Fig. 6**).
- A preponderance of cells that responded to both the CS^+ and US before training decreased their CS^+-evoked responses after training (**Extended Data Fig. 6a**).
- The set of cells with CS^+-evoked responses underwent substantial bi-directional plasticity during training, whereas the cells with CS^--evoked responses did not (**Fig. 2a,d**). That CS^--responsive cells generally have stable coding properties rules out the explanation that the

cells that depress their CS^+-evoked responses during training are simply those that are not reinforced by the US.

Moreover, the traditional Hebbian hypothesis requires, in its strictest form, precise temporal overlap between the neural input signals representing the CS^+ and US to induce potentiated neural responses to the CS^+ during learning[1]. In actuality, the amygdala can support associative fear conditioning without a hippocampal role provided the US follows the CS^+ within ~3 s (Ref. 2), and prior publications have cited this discrepancy as yet another weakness of the traditional Hebbian model[3]. This weakness of the Hebbian model is germane to the studies here, because we used an established form of short-trace fear conditioning[4-7] in which the CS^+ and US do not overlap and the US begins 800 ms after CS^+ offset.

This form of conditioning, which is hippocampal-independent[2] and amygdala-dependent[3-6,8-12] (**Extended Data Fig. 3**), allowed us to explicitly distinguish CS^+- and US-evoked activity, which was crucial for the analyses. As both classical delay and short-trace associative fear conditioning are strongly amygdala-dependent and have repeatedly yielded mutually consistent results regarding neural mechanisms and learned behavior[3], it is highly likely the BLA stores associative information via a temporally permissive plasticity rule requiring coincidence between the CS^+ and US to within ~3 s (Ref. 2). These results from past behavioral and neurophysiological studies further support our conclusions that BLA-dependent associative fear learning is not fully explained with a traditional, cellular Hebbian model of associative potentiation. Nonetheless, it remains likely that Hebbian plasticity contributes to the changes in the CS^+ representation. The relative prominence of Hebbian plasticity might depend on the exact time interval between the CS^+ and US presentations with the ~3 s permissive window, as

mechanisms that support Hebbian plasticity, such as NMDA-receptor-dependent potentiation, often require near-coincident, paired inputs[13].

The US-representation acts as a supervision signal that guides the transformation of the CS^+-representation.

The data here support the abstract interpretation that the ensemble representation of the US guides the transformation of the CS^+-representation to encode the learned association. The CS^+ population vector rotates directly towards the US, in the plane defined by the population vector representations of the US and that of the CS^+ prior to training (**Fig. 3f**). In this sense, the US representation is acting as a steering signal that governs the functional transformation of the CS^+ representation. This ensemble level model of supervised learning might also describe other forms of BLA-dependent associative learning, including those involving an appetitive US such as food reward or pheromone signals.

Can a tri-conditional cellular learning rule account for the BLA ensemble Ca^{2+} imaging data?

As noted above, a bi-conditional learning is insufficient to account for the diverse forms of plasticity that individual BLA neurons exhibit as the population vector representation of the CS^+ undergoes its re-scaling and rotation toward the US population vector (**Fig. 3f**). However, the addition of another component to the conventional Hebbian learning rule, such as the release of a neuromodulator or the activation of a widespread inhibitory circuit upon US presentation, can yield a tri-conditional rule capable of explaining the data.

With a tri-conditional learning rule, the outcome of neural plasticity depends on the presence or absence of three factors: signals encoding the CS^+ presentation; signals encoding the US presentation; and a global signal that permeates across the BLA circuitry, such as via neuromodulator release or a general inhibition of the BLA network, and that accompanies US

presentation. To illustrate, we present an example of a tri-conditional learning rule that is sufficient to account for the plasticity observed in the BLA ensemble neural calcium imaging data (**Supplementary Table 2**).

The top row of **Supplementary Table 2** accounts for the long-term stability of the CS^- representation, and the bottom two rows account for the bi-directional plasticity of neurons responsive to the CS^+. In contrast to the traditional Hebbian model's requirement for a strict temporal overlap between CS^+ and US to induce plasticity, here the global signal (such as a neuromodulator) that accompanies the US offers a potential mechanism[14,15] for extending the temporal window of plasticity induction to admit CS^+–US pairings in close temporal proximity but not strict concurrence. It is striking that past theoretical work has suggested networks performing supervised learning would need to augment the traditional Hebb rule with a tri-conditional plasticity rule[16], agreeing with our independent deductions from the BLA neural ensemble Ca^{2+} imaging data.

What accounts for the increase in CS^--evoked freezing after fear conditioning?

In addition to the bi-directional plasticity of cells' CS^+-evoked responses, we also found, to a lesser degree, bi-directional plasticity of the CS^--evoked responses (**Fig. 2d**) [9 ± 1% of cells were CS^--responsive before training *vs.* 11 ± 1% afterward; $P \leq 0.02$; Wilcoxon rank sum test]. The slight rise in the number of CS^--responsive cells after conditioning may explain the small rise in CS^--evoked freezing after training (**Figs. 1c, 2a**). Auditory inputs to BLA are less frequency selective than often seen within the auditory system[17], and the increased number of CS^--responsive cells could reflect a small fraction of inputs to BLA that transmit signals for both the CS^+ and the CS^- tones in the same axons. With such inputs, cells that potentiate their responses to the CS^+ would do so also for the CS^-. Alternatively, some cells might receive

separate synaptic inputs for the CS^+ and CS^- tones, but paired CS^+–US presentations might induce general changes, such as increased excitability of the soma or dendritic arbors, that heighten responses to the CS^-. Consistent with either of these scenarios, 32 ± 4% of cells that increased their responses to the CS^- also did so for the CS^+, ~3–4-fold more cells than would be predicted from the number responding to both tones before training (**Fig. 2a**). A third possibility is that the slight rise in CS^--evoked responses reflects a facet of conditioning that originates outside amygdala, such as in hippocampus[18-20], and leads to more freezing in response to any tones heard during conditioning.

References for Supplemental Note

1 Blair, H. T., Schafe, G. E., Bauer, E. P., Rodrigues, S. M. & LeDoux, J. E. Synaptic plasticity in the lateral amygdala: a cellular hypothesis of fear conditioning. *Learn Mem* **8**, 229-242 (2001).

2 Chowdhury, N., Quinn, J. J. & Fanselow, M. S. Dorsal hippocampus involvement in trace fear conditioning with long, but not short, trace intervals in mice. *Behav Neurosci* **119**, 1396-1402 (2005).

3 Sah, P., Westbrook, R. F. & Luthi, A. Fear conditioning and long-term potentiation in the amygdala: what really is the connection? *Annals of the New York Academy of Sciences* **1129**, 88-95 (2008).

4 Ciocchi, S. *et al.* Encoding of conditioned fear in central amygdala inhibitory circuits. *Nature* **468**, 277-282 (2010).

5 Herry, C. *et al.* Switching on and off fear by distinct neuronal circuits. *Nature* **454**, 600-606 (2008).

6 Senn, V. *et al.* Long-range connectivity defines behavioral specificity of amygdala neurons. *Neuron* **81**, 428-437 (2014).

7 Wolff, S. B. *et al.* Amygdala interneuron subtypes control fear learning through disinhibition. *Nature* **509**, 453-458 (2014).

8 Fadok, J. P., Dickerson, T. M. & Palmiter, R. D. Dopamine is necessary for cue-dependent fear conditioning. *J Neurosci* **29**, 11089-11097 (2009).

9 Gore, F. *et al.* Neural Representations of Unconditioned Stimuli in Basolateral Amygdala Mediate Innate and Learned Responses. *Cell* **162**, 134-145 (2015).

10 Rodrigues, S. M., Schafe, G. E. & LeDoux, J. E. Molecular mechanisms underlying emotional learning and memory in the lateral amygdala. *Neuron* **44**, 75-91 (2004).

11 Rosenkranz, J. A. & Grace, A. A. Dopamine-mediated modulation of odour-evoked amygdala potentials during pavlovian conditioning. *Nature* **417**, 282-287 (2002).

12 Tye, K. M. *et al.* Amygdala circuitry mediating reversible and bidirectional control of anxiety. *Nature* **471**, 358-362 (2011).

13 Rumpel, S., LeDoux, J., Zador, A. & Malinow, R. Postsynaptic receptor trafficking underlying a form of associative learning. *Science* **308**, 83-88 (2005).

14 Johansen, J. P., Cain, C. K., Ostroff, L. E. & LeDoux, J. E. Molecular mechanisms of fear learning and memory. *Cell* **147**, 509-524 (2011).

15 Johansen, J. P. *et al.* Hebbian and neuromodulatory mechanisms interact to trigger associative memory formation. *Proc Natl Acad Sci U S A* **111**, E5584-E5592 (2014).

16 Gutig, R. & Sompolinsky, H. The tempotron: a neuron that learns spike timing-based decisions. *Nat Neurosci* **9**, 420-428 (2006).

17 Bordi, F. & LeDoux, J. Sensory tuning beyond the sensory system: an initial analysis of auditory response properties of neurons in the lateral amygdaloid nucleus and overlying areas of the striatum. *J Neurosci* **12**, 2493-2503 (1992).

18 Phillips, R. G. & LeDoux, J. E. Differential contribution of amygdala and hippocampus to cued and contextual fear conditioning. *Behav Neurosci* **106**, 274-285 (1992).

19 McEchron, M. D., Bouwmeester, H., Tseng, W., Weiss, C. & Disterhoft, J. F. Hippocampectomy disrupts auditory trace fear conditioning and contextual fear conditioning in the rat. *Hippocampus* **8**, 638-646 (1998).

20 Quinn, J. J., Oommen, S. S., Morrison, G. E. & Fanselow, M. S. Post-training excitotoxic lesions of the dorsal hippocampus attenuate forward trace, backward trace, and delay fear conditioning in a temporally specific manner. *Hippocampus* **12**, 495-504 (2002).